

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Increasing returns to savings and wealth inequality

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1638462> since 2017-05-25T14:46:05Z

Published version:

DOI:10.1016/j.red.2007.02.003

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

Increasing Returns to Savings and Wealth Inequality*

Claudio Campanale[†]
Universidad de Alicante

November 8, 2005

Abstract

In this paper I present an explanation to the fact that in the data wealth is substantially more concentrated than income. Starting from the observation that the composition of households' portfolios changes towards a larger share of high-yield assets as the level of net worth increases, I first use data on historical asset returns and portfolio composition by wealth level to construct an empirical return function. I then augment an Overlapping Generation version of the standard neoclassical growth model with idiosyncratic labor income risk and missing insurance markets to allow for returns to savings to be increasing in the level of accumulated assets. The quantitative properties of the model are examined and show that an empirically plausible difference between the return faced by poor and wealthy agents is able to generate a substantial increase in wealth inequality compared to the basic model, enough to match the Gini index and all but the top 1 percentiles of the U.S. distribution of wealth.

Keywords: Wealth inequality, self-insurance, portfolio composition, increasing returns.

*I wish to thank Per Krusell and Rui Albuquerque for their suggestions during the development of this project. I also want to thank the editor Victor Ríos-Rull and two anonymous referees for useful comments. Financial support from the Instituto Valenciano de Investigaciones Económicas (IVIE) is gratefully acknowledged. All remaining errors and inconsistencies are mine.

[†]Departamento de Fundamentos del Análisis Económico, Universidad de Alicante, 03071, Alicante, Spain. Voice: ++34965903614 - Fax: ++34965903898. E-mail: claudio@merlin.fae.ua.es

1 Introduction

Empirical studies like Hurst, Luoh and Stafford (1998), Díaz-Giménez, Quadrini and Ríos-Rull (1997), Budría-Rodríguez et al. (2002) and Wolff (2000) have shown that earnings, income and wealth are very concentrated, with distributions that are skewed to the right. Of the three variables wealth is by far the most concentrated with a Gini coefficient of 0.78, while the same index for earnings and income is 0.63 and 0.57.¹ The latter fact is a regularity that is observed over time and across countries as well and has drawn considerable attention in the quantitative macroeconomic literature.

The basic framework used to explain this fact is the one in Aiyagari (1994) and Huggett (1996) and is based on a stochastic version of the standard neo-classical growth model featuring heterogeneous labor earnings shocks, missing insurance markets and borrowing constraints. Both models are successful at reproducing qualitatively the empirical evidence. However they are incapable of matching the data quantitatively so that various features, like heterogeneous subjective discount factors, bequest motives and entrepreneurship have been used in later work to improve the performance of the basic model.

Both the basic model and the extensions that followed share one key assumption about the assets available to the agents to carry out their saving plans. This assumption is that there is a single asset in the economy. A consequence is that all agents, no matter what their income or wealth is, face the same return on their investments. This assumption is clearly at odds with reality, since real world households may choose to hold assets as diverse in terms of return, risk and liquidity as for example housing and stocks or life insurance policies and checking accounts. To the extent that portfolio composition and returns vary systematically among households, these will have different incentives to save adding a further potential source of wealth inequality.

The goal of this research is to incorporate this basic feature of households' investment decisions in a stochastic, overlapping generation version of the neo-classical growth model and test whether the existence of increasing returns to savings is a quantitatively relevant source of wealth inequality. It turns out that it is: the empirically observed difference in the return faced by poor and wealthy households in the economy is sufficient to match the Gini index and the share of almost all percentiles of the U.S. distribution of wealth. This suggests that so far an important piece of the explanation for the massive concentration of wealth of real economies has been overlooked.

The model assumes exogenously that returns to saving are increasing in the

¹The values reported here are taken from Díaz-Giménez, Quadrini and Ríos-Rull (1997) and are based on the 1992 Survey of Consumer Finances.

level of wealth without modeling explicitly household portfolio choice; however, this feature of investment opportunities has strong support in the data. Empirical research by Bertaut and Starr-McLuer (2000), Kennickell et al. (2000) and Samwick (2000) clearly shows that the composition of households' portfolios shifts towards larger fractions of high-yield assets, like stocks and business equity, as the household's net worth increases. In the paper I first use Survey of Consumer Finance data on household balance sheets and data on returns of broad categories of assets from a variety of sources to give a precise characterization of the empirical relation between wealth and returns. This exercise reveals that while the poorest 60 percent of the population faces an average return to its wealth which is close to 1 percent, the top 1 percent invests its wealth at an average return between 4.5 and 6 percent. Then I interpolate the empirical schedule and use it in a standard model with uninsurable idiosyncratic earnings risk that mixes the life-cycle and dynastic framework. The properties of the resulting stationary distribution of wealth in the two cases of constant and increasing returns to savings are compared revealing that when the estimated return function is used a substantial boost to wealth inequality follows closing the gap between the quantitative prediction of a standard model with constant returns and the data.

Before moving to the remaining sections of the paper it is important to spend a few words about the interpretation of the positive relation between net worth and portfolio returns that is observed in the data and, based on that evidence, is assumed here. There are two alternative but not mutually exclusive stories that can be told to explain this relation. The first is a market imperfection one: it may be necessary to pay information costs to gain knowledge of the functioning of some asset markets and even then other trading costs are required to actually participate in those markets. As a consequence only households that have accumulated enough wealth may find it attractive to spend the time and money needed to participate in those higher return markets. In support of this view come a number of studies like, for example, Paiella (2001) and Vissing-Jørgensen (2001) about the costs of participating in the stock market and by Hong, Kubik and Stein (2001) that find participation to the stock market being positively related to sociability as a result of the effect that communication with peers has in lowering information costs. The second story is a behavioral one: according to this interpretation some agents dislike some assets and decide not to participate even if this would be optimal based on their risk preferences and on the asset return. In support of this possibility the paper cited above by Hong et al. (2001) reports that participation in the stock market in the U.S. is substantially higher for white than for non-white even after controlling for wealth, income, education and survey measures of risk tolerance; Guiso, Sapienza and Zingales (2004)

reach a similar conclusion with Italian data when comparing participation rates between southern and northern Italians.²

The model presented in this paper is more consistent with the first interpretation since it implies a positive feed-back from wealth accumulation to higher returns and again to further accumulation; moreover it does not assume heterogeneity in tastes. However it is not in contrast with the behavioral story: by showing that small differences in the return to assets can generate a realistic concentration of wealth it says that small ex-ante differences in behavior concerning portfolio composition may lead to the large observed wealth inequality. Moreover the same positive relation between net worth and the return on savings assumed in this paper would still be obtained as an ex-post result in a behavioral one.

The rest of the paper is organized as follows. In Section 2 I first review the quantitative literature on wealth inequality. In Section 3 I present an account of the empirical evidence on wealth inequality and on portfolio composition at different wealth levels; I then construct an empirical schedule that maps net worth into average portfolio returns. In Section 4 I describe the model, in Section 5 the choice of parameters and in Section 6 the results. Finally Section 7 concludes. Details about the construction of the data used in Section 3 and the numerical solution of the model are given in Appendix A and B respectively.

2 Literature Review

A large number of papers present quantitative models that attempt to explain the observed wealth distribution. These models share a set of basic assumptions. First they are populated by agents who receive an exogenous stochastic flow of income. Second, markets are assumed to be incomplete so that it is not possible to fully insure consumption risk. Finally there is some form of borrowing constraints. A notable example of this kind of models is Aiyagari (1994). Agents use accumulated savings in order to buffer negative shocks to income and therefore to smooth consumption. While agents are ex-ante homogeneous, ex-post each of them will have experienced a different history of past realized incomes leading to a different level of accumulated wealth. This model generates a distribution of wealth that is more concentrated than the distribution of income, a feature that is qualitatively consistent with the data. However at a quantitative level it grossly under-predicts the observed concentration of wealth.

²The three authors suggest that financial contracts and stocks in particular are trust-intensive contracts. They then exploit variation within Italy of measures of social capital (of which trust is an important element) and show that these are positively related to the use of checks, participation in the stock market and the availability of credit.

In particular it fails to explain the two tails of the distribution, that is, the very low level of wealth accumulation by poor agents and the accumulation of huge estates at the very top of the wealth distribution.

The model in Aiyagari (1994) considers an economy populated by infinitely-lived dynasties in which saving occurs for precautionary reasons. In a related paper Huggett (1996) uses similar assumptions about market structure, but casts the model in a finite-horizon framework where agents face a realistic lifetime profile of earnings and go through the working and retirement stages of life. In this framework saving also occurs to finance retirement consumption. The model fares well in terms of matching the Gini index but it obtains this result by having a large fraction of households with no or even negative wealth while still underestimating the large accumulation of assets of the very rich. Moreover those with no wealth are entirely concentrated among very young households that face an upward sloping earnings profile and would like to borrow.

Following these two papers various mechanisms have been proposed to improve the ability of quantitative models to match the observed concentration of wealth. These attempts may be broadly classified based on whether their main focus is on the left or the right tail of the distribution.

A prototypical example of the first group is the paper by Hubbard, Skinner and Zeldes (1995). Their model is cast in a finite-horizon framework and features both earnings and health risk. The crucial element is the presence of means-tested government programs that provide a safety net in the form of a floor on consumption in case of very bad luck. This induces very poor agents not to accumulate assets at all and rely on public social insurance instead. While not directly focused on measuring wealth inequality this model is able to generate a substantial number of agents with very low or no savings at all without having them entirely concentrated among younger agents.

Another institutional feature that has the potential to reconcile the data on the uneven wealth distribution with the output of quantitative models is a progressive social security system. This has been used by Domeij and Klein (2002) to account for the large portion of Swedish households with very little wealth and, coupled with lifetime differences in earning abilities, has also been proposed by Huggett and Ventura (2000) to explain why low income households as a group save a lower fraction of their income than high income households do.

As far as the right tail of the wealth distribution is concerned two mechanisms have been proposed so far. Based on the empirical observation, reported in Gentry and Hubbard (2000), that entrepreneurs both make a significant share of very wealthy households and tend to have higher wealth-income ratios, Quadrini (2000) constructs a model where entrepreneurship is recognized as the critical

element to add to a quantitative model to generate a realistic wealth concentration. In his model imperfections in financial markets drive a wedge between borrowing and lending rates so that the marginal return to saving and investing in the private firm is higher than market returns. Moreover consistent with the empirical evidence, Quadrini assumes that the income flow generated by a business is more risky than the income of paid employees, increasing precautionary saving. The model is then able to generate a more realistic wealth concentration and to account for the higher wealth-income ratio and upward mobility of entrepreneurial households.

The second mechanism exploits intergenerational links in the form of altruism and correlation between the earning abilities of successive members of a family. Castañeda et al. (2002) is an example along these lines: the authors exploit intergenerational links in a model with endogenous labor supply and a stylized representation of the U.S. progressive taxation system to check if it is possible to find a process for earnings that allows the model to match the distribution of earnings and wealth simultaneously. They find that the answer to this question is positive even though it comes at the cost of an earnings process with some unusual features. In a slightly different fashion De Nardi (2001) also constructs a model populated by finitely lived agents in which parents and children are linked by voluntary bequests and persistence within families in earnings abilities. Her model is calibrated on U.S. and Swedish data and shows how the two intergenerational links are important to explain the emergence of the large estates we observe at the top of the wealth distribution.³

Finally a completely different approach has been followed by Krusell and Smith (1998). The key feature of their model is the assumption of heterogeneous subjective discount factors. While many economists would look with suspicion at a model based on an unobservable variable like the discount factor, there is some experimental evidence in favor of preference heterogeneity.⁴ The economy in Krusell and Smith is populated by infinitely lived agents whose discount factor changes stochastically over time with an average frequency close to the average length of life. The consequence of this assumption is that some agents will be patient, accumulate wealth and therefore fix the equilibrium interest rate; the

³The two approaches are brought together by Cagetti and De Nardi (2002). First, they explicitly model a market friction that limits entrepreneurial borrowing generating higher returns to investment in own firms than on the market. Second, their economy is populated by stochastically aging agents who go through the stages of working life, retirement and death, therefore allowing for voluntary bequests. The joint operation of higher marginal returns to business investment and the bequest motive enable their model to reproduce the high concentration of wealth at the top of the distribution, although it is still true that the very wealthy are active or retired entrepreneurs, which leaves the empirically observed share of wealthy non entrepreneurs unaccounted for.

⁴See for example Barsky et al. (1997).

rest will have a discount factor well below the interest rate and therefore act as hand-to-mouth consumers. As a result the model is able to generate both a large number of agents with low or no assets at all and an empirically plausible concentration of wealth at the top of the distribution.

3 The Empirical Evidence

The purpose of this section is to describe in details the composition of household portfolios along the wealth distribution and characterize the relation between asset holdings and returns. The main finding is that richer households tend to have more complex portfolio structures with a larger fraction of their net worth held in high yielding assets than poorer households. Using data on portfolio composition by level of net worth, together with data on returns to different assets I then compute an empirical return schedule. It will be shown that the range of this schedule is not large but still not negligible.

3.1 Wealth and Portfolio Composition

In this subsection I report a detailed analysis of the changes in households' portfolios with the level of net worth. There are some good surveys on the topic like Bertaut and Starr-McCluer (2000), Heaton and Lucas (2000), Kennickell et al. (2000) and Samwick (2000). The main messages that consistently emerge from all of these studies is that the structure of family portfolios increases in complexity as their wealth increases and that wealthier households invest larger shares of their savings in higher-return and higher-risk assets. The analysis presented here confirms those findings. Its main distinctive feature is that since it is aimed at characterizing an empirical return function that will then be used in the quantitative model presented in the next sections, it is based on a partition of the wealth distribution that is finer than what is commonly used in the literature. A summary of the results of this analysis is presented in Table 1 and 2; the source of data used here is the 1998 issue of the SCF.⁵ In Table 1 I report the percentage of households who own the particular asset indicated in the first column of the table. This is done for a subset of assets, that is, liquid accounts, stock, primary residence and business equity. Two suggestions come from this table. First, the ownership of each of the four assets increases with wealth confirming the finding that richer households own more complex portfolios. Second the pattern of ownership along the wealth distribution is quite

⁵The same analysis was performed on other issues of the SCF. Results are very similar confirming the relative stability of patterns of portfolio shares across the wealth distribution found in other studies. These results are not reported here in order not to burden the text with too many large tables, however they are available from the author upon request.

Table 1: Asset Ownership by Net Worth Percentiles

	Net Worth Percentiles						
	0-40	40-80	80-90	90-95	95-98	98-99	99-100
Liquid	79.47	97.51	99.91	100.00	99.95	99.80	99.88
Stocks	25.83	55.31	77.49	86.91	91.60	87.76	97.35
Home	30.88	87.43	95.06	93.54	93.58	97.73	96.24
Business	2.75	10.65	18.76	33.08	37.41	56.53	62.24

different across different assets. It is clear from the table that liquid accounts, a group of assets that pay a very low return, are very widespread even among the bottom group of the distribution and their use becomes universal starting from the 40 to 80 percentile group. A similar pattern is observed for the other low paying asset, that is, housing, with the exception that the ownership rate is much lower in the bottom group, likely because of the indivisibility of the initial down-payment needed to buy a house. On the contrary stock,⁶ a high return asset, is owned by only a quarter of households in the bottom 40 percent of the distribution and still by only a half of families in the 40 to 80 group; it becomes almost universal only starting from the 95 to 98 percentile group. This pattern of substantial increase is even more striking when we consider the other high-yield asset, that is, private equity: here only about 3 percent of households in the bottom 40 percent of the distribution owns the asset, a percentage that climbs up to 63 in the top 1 percent.

While Table 1 reports data on participation to different asset markets, Table 2 takes an alternative perspective and looks at the shares that assets with different returns represent in the average portfolio of families belonging to different percentiles of the wealth distribution. It does so for the complete set of possible assets grouped into broader groups: liquid accounts, bonds, stocks, home, investment real estate and business assets plus two residual categories that include financial and nonfinancial assets not otherwise classified. Among financial assets, liquid accounts and bonds, show a hump-shaped profile but the variation is minor along the whole wealth distribution: the combined share is about 8.3 percent in the bottom 40 group and it is still only 11.2 percent in the top 1 percent of the distribution with a slightly higher peak in the middle. Stock instead shows a dramatic rise from forming only 6 percent of assets in the bottom 40 group to about 27 percent in the 95 to 98 percentile of the distribution after which it stabilizes. Moving to nonfinancial assets, we see that the share of primary residence in total assets declines monotonically and by a large

⁶Here the definition of stock includes any form of ownership from directly held stocks to stock held through mutual funds and retirement accounts. See Appendix A for details about data construction.

Table 2: Portfolio Composition by Net Worth Percentiles

	Net Worth Percentiles						
	0-40	40-80	80-90	90-95	95-98	98-99	99-100
<i>Financial Assets</i>							
Liquid	5.38	6.89	7.15	7.90	8.07	6.08	4.31
Bonds	2.94	5.64	7.22	7.50	8.28	8.05	6.85
Stock	6.0	10.71	17.29	20.15	26.63	25.42	26.42
Other Financial	8.11	9.42	12.03	13.28	9.76	9.16	7.94
<i>Non Financial Assets</i>							
Home	53.00	51.49	36.81	26.26	20.75	14.08	7.67
Real Estate	3.96	5.05	9.25	13.28	9.97	18.74	11.21
Business	0.96	2.74	5.07	7.73	13.06	16.60	32.34
Other Nonfinancial	19.67	8.07	5.19	3.91	3.47	1.86	3.18

amount from 53.0 percent in the bottom 40 group, down to about 8 percent in the top 1 percent of the wealth distribution. A similar dramatic change, but in the opposite direction can be observed in the pattern of business equity ownership: this asset represents less than one percent in the average portfolio of the 40 percent poorest segment of the population and rises to above 30 percent in the top 1 percent segment. If we further aggregate assets into broader groups we see that the combined share of stock and business equity, the two high-yield assets moves up from only 7 percent of wealth in the bottom 40 group to about 28 percent in the 90 to 95 percentile and up to almost 60 percent in the top percentile of the distribution. Summarizing, Table 2 clearly shows that as wealth increases the share of high-yield assets increases and that of low-yield assets decreases, so that households face a return schedule that is increasing in their asset holdings. This statement will be made more precise in the next section where an empirical return schedule is constructed.

3.2 Portfolio Returns

In this section I take the evidence on portfolio composition by percentiles of the wealth distribution described in the previous section and using data about the return to different categories of assets I construct an empirical return function that maps wealth holdings into returns on that wealth. The goal of this section is to give a precise characterization of the relationship between returns and asset holdings that can be used in the modeling section of the paper. For this reason I use a partition of the wealth distribution that is finer than the one that appears in Tables 1 and 2. In particular I divide the population into deciles up to the eightieth percentile, into 5 percent groups from the eightieth to the ninety-fifth and then I consider separately the ninety-five to ninety-eight percentile

group and the ninety-eight to ninety-nine; finally I divide the top 1 percent of the distribution into two half percentiles, for a total of fifteen observations. For each of these groups I compute total asset holdings and the average share of each category of assets. In practice I eliminate from the computation the assets in the residual categories since it is difficult to impute a measure of their return. The share of the group labeled “other financial assets” shows a hump-shaped pattern over the wealth distribution with very modest variations, so that its exclusion does not affect the computed average returns differently for different wealth groups. The share of the group labeled “other nonfinancial assets” instead shows a strong declining trend with the bottom 40 percent of households holding about 20 percent of their wealth and the top 1 percent holding only 3 percent of theirs in these types of assets. Notice though that a large part of this category is made by vehicles, a durable good that clearly has negative returns as an asset, so that inclusion of this group would actually further increase the magnitude of increasing returns to savings. The other adjustment that I make is to consider primary residence and investment real estate as a single asset since the latter includes a large proportion of housing as well. By doing this regrouping I end up splitting wealth into five classes: liquid accounts, stocks, bonds, property and business assets.

Interest-bearing accounts are a heterogeneous group of assets that includes checking and saving accounts, which pay a negative real interest, and certificates of deposit and money market accounts, which pay a small positive interest; therefore, I conventionally set the return to this category to 0. I set the return to stocks to 8 percent, the value reported in Jagannathan et al. (2000) for the return to the S&P 500 index for the period 1926 to 1999. Based on Moskowitz and Vissing-Jørgensen’s (2002) claim that the return to private equity is no different than the return to the public equity index I also set the return to business assets at 8 percent. As far as bonds are concerned, these are again a heterogeneous category of assets including government, corporate and foreign bonds as well as municipal and state bonds that have a preferential tax treatment. Not having a finer subdivision of the category I attribute to it the return to 20-year U.S. Treasury securities of 1.9 percent reported in Jagannathan et al. (2000) and referring again to the period 1926-1999. Finally, Goetzmann and Spiegel (2000) report that according to the Office of Housing Enterprise Oversight the real price of housing has increased at a 0.5 percent annual compounded rate over the period 1980 to 2000. While the return to residential property includes the housing services that it provides, this class of assets has special costs like property taxes and maintenance. Moreover high costs and risks of transaction may have a strong negative impact on the return, especially when the holding period is short. The two authors then suggest that the return to this asset may be even

Table 3: Returns by Net Worth Percentiles

Percentiles	Normalized Assets	Percentage Return
100–90	-0.028	1.99
90–80	0.009	0.86
80–70	0.038	1.11
70–60	0.100	1.13
60–50	0.195	1.31
50–40	0.317	1.55
40–30	0.483	1.74
30–20	0.751	2.10
20–15	1.105	2.42
15–10	1.474	2.74
10–5	2.275	3.10
5–2	4.463	4.02
2–1	9.795	4.14
1–0.5	16.256	4.79
top 0.5	51.438	5.80

lower than the 0.5 percent per year that their price appreciation suggests. Given all these considerations I take the value of 0.5 percent as the return to property. With these numbers I can construct the return to the average portfolio of the different percentiles of the wealth distribution mentioned above; the results are reported in Table 3. The table reports in the second column the average asset holdings in the quantile indicated in the first column normalized by the population average and in the third column the return to that wealth⁷. There are two points that come out of the table. First, the average return that households face is monotonically increasing in the level of net worth except for the bottom two deciles. However this does not contradict the general observation since the sample composition at the very bottom decile turns out to be anomalous compared to the whole distribution. If one looks at the details this group is made by two very different types of households: very few ones have very large assets and even larger debts mostly related to entrepreneurial activity, while the vast majority has no or very little asset holdings. Consequently the composition of asset holdings in this group will be dominated by the large holdings of private equity of a handful of households in the decile biasing upwards the computed return. The second is the magnitude of the difference between the return faced at the top and at the bottom of the distribution. In 1998 the average household in the top 0.5 percent wealthiest households held about 51 times average net

⁷Notice that households are classified into net worth percentiles as it is standard practice in the empirical portfolio literature. The return schedule instead maps asset holdings into portfolio yields to make it consistent with no-borrowing constraint in the modeling section

worth and faced a return on its assets of 5.8 percent, while the average return faced by a household in the 90 to 80 percentile of the distribution, which held a puny 1 thousandth of average wealth, was 0.86 percent. The difference is then of about 5 percentage points a figure which is non negligible.

4 The Model

The model economy studied in this article is based on the neoclassical growth model with uninsured idiosyncratic risk and no aggregate uncertainty. It is populated by a large number of households with identical preferences and finite life that are linked to form infinitely lived dynasties. Each household faces an age changing probability of survival and goes through the stages of working life and retirement; after death it is replaced by a newly born household that inherits its financial wealth and part of its earning ability. The shock to permanent earning ability is household specific and uninsurable. During working life all households face a common persistent stochastic process that co-determines their period productivity and whose realizations are idiosyncratic and uninsurable. In the experiments where a bequest motive is active this is of the altruistic form. In the subsections that follow I describe formally the features of this economy. Since the interest is on steady states time indexes are omitted and the index t is reserved to denote agents' age.

4.1 Demographics

The economy is populated by a continuum of households with finite life. Time is discrete and each period corresponds to one year. The maximum possible life length is 100 years. Households enter the model at real-life age 20 as workers. After that they face an age dependent probability of survival that I indicate with p_t ; if they survive long enough they retire at age R . Consequently the value of T and R are 80 and 45 respectively. After death the household is replaced by a 20 year old descendant who inherits its financial wealth, if any, and part of its permanent productivity.

4.2 Earnings and Pensions

During working life agents supply inelastically the amount of efficiency units of labor they are endowed with in exchange for a wage; after age 65 they retire and receive a pension benefit from the government. Earnings are the product of three components. First there is a deterministic component that is common to all households in the economy and is meant to capture the hump in average life cycle earnings observed in the data; I will denote this component with $H(t)$. Second

there is a permanent component, indicated with θ , that captures differences in earnings ability that are household specific and fixed in the course of life. This component is assumed to follow an $AR(1)$ process in logarithms, so that part of the parent's earnings ability is inherited by its descendant. Finally there is a persistent component, denoted with z , that also follows an $AR(1)$ process in logarithms and that determines the household specific yearly evolution of productivity. Summarizing if we denote with y_t the amount of efficiency units of labor available to the household at age t then they will be equal to:

$$y_t = H(t)\theta z_t \quad (1)$$

and

$$\log \theta = \rho \log \theta_{-1} + \varepsilon_\theta \quad (2)$$

$$\log z_t = \varrho \log z_{t-1} + \varepsilon_z \quad (3)$$

where ε_θ and ε_z are i.i.d. normally distributed random variables. The parameter ρ measures the degree of intergenerational correlation of earnings and ϱ determines the degree of persistence of earnings during an agent's working life. Once past retirement age R the persistent component of earnings z_t is conventionally set to 0 and pension income substitutes wage earnings. Pension benefits are constant during retirement and consist of two components: a fixed part \bar{b} and a variable component that depends on the agent's permanent component of earning ability. If we denote this term with $b^v(\theta)$ we can write the overall transfer income received by a retired agent as:

$$b(\theta) = \bar{b} + b^v(\theta) \quad (4)$$

4.3 Preferences

Households do not value leisure so that period utility is defined by a utility index $u(c_t)$ where u is strictly increasing, strictly concave and satisfies the standard Inada condition for interior solutions. They are altruistic, discount future own utility at rate β and apply a further discount factor $0 \leq \gamma \leq 1$ to the descendent's utility. In this way the model captures the life-cycle model — $\gamma = 0$ — and the fully altruistic model — $\gamma = 1$ — as the two polar extremes on a continuum.

4.4 Technology

Output Y_t is produced using the aggregate available capital K_t and the aggregate supply of labor L_t which is normalized to 1 for convenience. Households own the capital and rent it to firms. Capital depreciates at a variable rate. More specifically there exists a fixed component of depreciation that is common to

all capital and is denoted by δ . In addition to it there is an individual specific component that depends on the amount held by the household. This latter component will be denoted with $\phi(k)$ where ϕ is a strictly decreasing function with $\lim_{k \rightarrow +\infty} \phi(a) = 0$.

4.5 Government

The government in the model economy studied taxes household income and estates and uses the revenues from taxation for consumption and to make transfers to retired households. I assume that there are three types of taxes: a proportional income tax, an estate tax and a payroll tax, denoted respectively by τ , τ_e and τ_s . The proceeds from the income and estate taxes are used to finance government consumption G and a government budget balance restriction applies in every period. The payroll tax is separately collected and used to finance the transfers made to retired households; it is assumed that the social security system is balanced every period.

4.6 The Household Decision Problem

The household state variables are its asset holdings and its shock to permanent and persistent components of efficiency units of labor, that is, the triple (a_t, θ, z_t) . notice that while asset holdings and the persistent component of individual productivity change as an agent ages so that they need to be indexed by t , θ is fixed for a given household, therefore it is not indexed. Aggregate state variables, that is, the measure of agents over individual states are part of a complete description of an agent's state variables. Here the focus of the analysis is on steady states only so that this measure can be treated as a parameter and omitted. Given the finite horizon faced by each single household and the periodic component of its earnings and pension income, age must be added as a further state. The household dynamic programming problem then reads:

$$V_t(a_t, \theta, z_t) = \max_{c_t, a_{t+1}} u(c_t) + \beta E\tilde{V}_{t+1}(a', \theta', z') \quad (5)$$

$$E\tilde{V}_{t+1}(a', \theta', z') = p_{t+1}EV_{t+1}(a_{t+1}, \theta, z_{t+1}) + \gamma(1 - p_{t+1})EV_1(a_1, \theta, z_1) \quad (6)$$

$$a_{t+1} = \tilde{a}_{t+1} \quad (7)$$

$$a_1 = \tilde{a}_{t+1}(1 - \tau_e(\tilde{a}_{t+1})) \quad (8)$$

The maximization is performed subject to the following restrictions:

$$c_t + \tilde{a}_{t+1} \leq a_t(1 + r(a_t)(1 - \tau)) + I_R(1 - \tau - \tau_s)wy_t + (1 - I_R)b(\theta) \quad (9)$$

$$c_t \geq 0, \quad \tilde{a}_{t+1} \geq 0 \quad (10)$$

and equations (1),(2),(3). Equation (5) states that the indirect utility of an age t agent is the sum of the utility it derives from current consumption plus the discounted expected continuation utility. In turn expected continuation utility $E\tilde{V}_{t+1}(a', \theta', z')$ is the sum of two terms. First, if the agent survives, which happens with probability p_{t+1} , it will turn age $t+1$, have the full amount of assets $a_{t+1} = \tilde{a}_{t+1}$ chosen the previous period, keep the same realization of permanent earnings ability θ and get a new draw of the i.i.d. shock leading the permanent component of productivity to the value z_{t+1} . Second, if the agent dies, which happens with probability $1 - p_{t+1}$ it will be replaced by an age 1 agent; this new agent will inherit its wealth minus the estate tax — $a_1 = \tilde{a}_{t+1}(1 - \tau_e)$ —, get a new draw of the lifelong component of the endowment of efficiency units of labor, that is, θ_+ and a draw from the first year distribution of persistent productivity z_1 . In this case the indirect utility, denoted $V_1(a_1, \theta_+, z_1)$, is further discounted at the rate γ allowing in this way less than perfect altruism. The budget constraint (9) describes the sources and uses of funds available to the agent. In this equation I_R is an indicator function that takes a value of one if the agent is working and 0 if it is retired. In the former case it receives earnings wy_t , the product of the wage rate times the endowment of efficiency units of labor, net of the income and social security tax; in the latter case it receives the pension benefit $b(\theta)$. Beside wage and transfer income the household receives income from its asset holdings: here $r(a_t)$ describes the assumed dependence of return on the amount of asset holdings. With a slight abuse of notation I denote with r the marginal product of capital gross of depreciation, so that $r(a_t) = r - \delta - \phi(a_t)$. Finally τ is the constant income tax rate.

4.7 Equilibrium

In order to simplify the notation the letter s will be used to summarize the individual state variables including its age, that is, $s \equiv (a, \theta, z, t)$ and age subscripts are dropped. Also, let x be a stationary measure of households. A stationary equilibrium for this economy is a value function $V(s)$, household decision rules $\{c(s), a'(s)\}$, government policy $\{\tau, \tau_e(a'(s)), \tau_s, b(s), G\}$, factor prices (r, w) , macroeconomic aggregates $\{K, L, T, T_s\}$ and the stationary measure of households x , such that:

1. Total factor inputs, tax revenues and transfer payments are obtained aggregating over households:

$$\begin{aligned}
- K &= \int a dx \\
- L &= \int y(s) dx \\
- T &= \int \tau(ar(a) + wy(s)) dx + \int p(s)\tau_e(a'(s)) dx
\end{aligned}$$

$$- T_s = \int \tau_s w y(s) dx$$

where $p(s)$ denotes the probability that a state s agent dies, that depends on its age only.

2. Given prices, taxes and transfers $V(s)$ is the optimal value function and $\{c(s), a'(s)\}$ are the associated decision rules.

3. Prices are determined competitively, that is,

$$r = F_1(K, L) \quad \text{and} \quad w = F_2(K, L)$$

4. The goods market clears:

$$\int [c(s) + a'(s)] dx + G = F(K, L) + (1 - \delta)K - \Delta$$

where Δ is the total amount of the variable depreciation, that is, $\Delta = \int a\phi(a) dx$

5. The government and social security administration budget are satisfied:

$$- T = G$$

$$- T_s = \int b(s) dx$$

where $b(s)$ is the benefit received by a state s agent, which depends on his age and permanent ability and the integral on the right hand side gives the total social security expenditures.

6. The measure of households is stationary, that is:

$$x(B) = \int Q(s, B) dx$$

Here $Q(s, B)$ is a transition function that gives the probability that an agent that is in state s in the current period will have state $s' \in B$ in the following period. The transition function is defined by the joint operation of the agent optimal decision rules and the exogenous stochastic processes for age and labor efficiency.

4.8 Discussion

In this section I present a discussion of some of the assumptions made. First of all the characterizing feature of the present research is the assumption that capital depreciates at different rates depending on the amount accumulated by the household owning it. This assumption captures the fact, reported in the empirical section above, that households invest larger shares of their wealth in higher return assets as they become richer. The use of a reduced form instead of an explicit model of portfolio choice has two motivations. The first and most important one is that currently the issue of portfolio choice is largely unresolved, in particular as far as the main point of this paper is concerned, that is, the

positive relation between wealth and returns.⁸ Second, introducing a model of portfolio choice that is sufficiently accurate would make the solution of the steady state general equilibrium computationally unfeasible. Despite the simplification this reduced form introduces, it still has some adverse consequences on the computational burden of the solution that put constraints on other modeling choices. Once returns are increasing in wealth, the individual budget set is not any more convex, consequently the value function is not concave and its marginal value not monotonically decreasing. This gives rise to the possibility of multiple local maxima making it impossible to use fast algorithms like Newton, Brent or even bisection methods to find the optimal asset choice given the current state variables.⁹ A global maximization routine is then needed and in this specific case a direct search method was used; more details about its features are given in Appendix B. One consequence of the need to adopt a global optimization algorithm is that it slows down the overall numerical solution constraining some other choices as discussed below. The demographic structure of the present model economy borrows from Castañeda et al. (2002) and Cagetti and De Nardi (2002) in assuming that adjacent households within a dynasty replace each other rather than overlapping. It blends it with a traditional life-cycle framework where agents face age changing death probabilities and a hump shaped profile of earnings followed by retirement. This framework was chosen to balance two different goals. On the one hand the desire to give a careful representation of inequalities in asset holdings arising from life-cycle determinants of saving.¹⁰ On the other the need to avoid the excessive complications that would arise in a model with intergenerational altruism if consecutive members of dynasties overlapped; in this case in fact the descendants' state variables would enter the parent's problem and vice-versa increasing substantially the dimension of the state space.¹¹ Similar considerations are behind the choice of having progressive social security benefits linked to permanent earnings differences. The choice, common in the literature, of assigning a fixed pension benefit to all retired agents would tend to magnify the relative savings of high-earners

⁸The interested reader can find in the book edited by Guiso et al. (2002) a detailed survey about the methodology, data and theory as well as a large number of references.

⁹All these methods require that the objective function be single-peaked and Newton method puts even stronger restrictions since it needs a concave objective. See Brent (1973) for details.

¹⁰The only minor drawback of this formulation is that since every household inherits at the beginning of life, very young agents that would otherwise be at their borrowing constraint may instead have positive amounts of wealth, leading to some underestimation of actual wealth inequality. The problem though is likely to be minor given that a large number of inheritances in the model are small.

¹¹De Nardi (2004) partially avoids this problem by assuming "warm glow" altruism; however even under this assumption she needs one more round of computation because at the very least one needs to assume that the descendant makes some forecast about his parent's wealth that are consistent with the actual population distribution

compared to low earners and therefore wealth inequality. This is because if one sets the replacement ratio equal to its population average for everybody in the economy, the replacement ratio of high ability types would be lower and that of low ability types would be higher than it is in reality magnifying savings for retirement of the former group compared to the latter. The choice made here mitigates this problem without the need to add a further state variable to keep track of cumulated past earnings, thus saving the associated computational burden.

5 Calibration

In this section I describe the way in which parameters are chosen: some of them are taken from other studies while the rest are set so that the model generated data match some target ratios taken from the US economy. The description of technology requires specifying a functional form for the production function and the variable part of capital depreciation and assigning values to the parameters defining them. The production function is assumed to be of the Cobb-Douglass form: $Y = AK^\alpha L^{(1-\alpha)}$. The share of income that goes to capital α is set to 0.36 a value taken from Cooley and Prescott (1995) and the constant A is used as a normalization factor to make the wage rate in the economy equal to one. The fixed portion of depreciation δ is set to 0.08 the value used by Aiyagari (1994). Given a target value for the capital output ratio of three this implies that the maximum interest rate net of depreciation and gross of taxes is 4 percent. This value is — approximately — enjoyed only by very wealthy households; because of the variable component of depreciation poorer households will face lower returns. In choosing a functional form for the variable depreciation function $\phi(a)$ it is taken into account that a visual inspection of the empirical return function obtained from SCF data suggests that returns are clearly monotonically increasing in asset holdings and that they converge asymptotically. Two alternative functional forms were considered. The first one is a logistic, that is:

$$\phi(a) = \frac{c_1}{1 + c_2 \exp(a - c_3)}. \quad (11)$$

This functional form allows for the possibility of a non concave relation between wealth and returns that might not be detectable from an inspection of the data. The second alternative is a hyperbolic function that forces concavity, that is:

$$\phi(a) = \frac{c_1}{a^{c_2} + c_3}. \quad (12)$$

A graphical representation of the approximation obtained with these two functions is postponed to the result section; here I describe the procedure to fix the

Table 4: Model Parameters

α	0.36
δ	0.08
σ	2.
ϱ	0.5
$std(\varepsilon_\theta)$	0.75
ρ	0.93
$std(\varepsilon_z)$	0.158
G/Y	0.192
τ_e (exemption)	15.0
bv/Y	0.1242
τ_s	0.143

parameters of these functions. Using SCF data the population was split into 15 net worth groups. The average weighted portfolio composition and average asset holdings was obtained for each of these groups and then, using historical data on asset returns, the return on this average portfolio was computed. In this way I obtained an empirical schedule that maps average asset holdings into average returns. This schedule was then interpolated with the two functions specified above and the parameters were found by minimizing the sum of the squared distances between theoretical and empirical values subject to the constraint that c_1, c_2 and c_3 be non- negative. Preferences are described by a standard power utility function, that is, $u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma}$ and σ is set equal to 2, a value well within the range usually used by macroeconomists —see for example Aiyagari (1994). In addition I need to fix the subjective discount factor on own and descendant utility. These two parameters are determined endogenously. The discount factor on own utility β is set to clear the market when the capital-output ratio is 3. The discount factor further applied to the descendant’s utility γ is determined so that the median bequest to average earnings ratio in the models with increasing returns matches the value found in the data. The value for earnings are taken from Budria et al. (2002) and the value for median bequest is the one reported by Hurd and Smith (1999) for single decedents. Two comments are needed on this choice: first I prefer to use the bequest left by single because in general a surviving spouse inherits at the first death in the couple so that the bequest left by the decedent in a couple would overestimate the actual amount of intergenerational transfer of wealth; second the median value of bequests was preferred to the average one because the data source used by Hurd and Smith — i.e. the Asset and Health Dynamics among the Oldest Old or AHEAD — is a small random sample and it is known that this makes the measurement of variables that are very concentrated like bequests unreliable because of the poor

representation of the top of the distribution.

The endowment of efficiency units of labor is determined by three components. The deterministic part $H(t)$ is obtained from the life cycle profiles estimated by Cocco, Gomes and Maenhout (2005) on high school graduates in the PSID. While this choice may seem restrictive, this profile is also consistent with the one estimated by Hansen (1993) on the overall population. Calibrating the random component of earnings requires specifying four parameters, that is, the autocorrelation coefficient and the variance of the processes for θ and z_t . Solon (1992) and Zimmermann (1992) estimate that the intergenerational correlation of earnings is bounded below by 0.4 and can be as high as 0.6 so I set the autocorrelation of the productivity inheritance process ϱ to 0.5. The variance of ε_z is taken to be 0.025 a value in the range of available estimates — see for example Hubbard, Skinner and Zeldes (1994). Once these parameters are chosen, the variance of ε_θ , the innovation in the inherited component of earning ability, determines uniquely the concentration of first year workers so it was picked to match the Gini coefficient of earnings in the age group below 25 computed by Budría-Rodríguez et al. (2002) using the 1998 issue of the SCF. Finally the autocorrelation coefficient of the persistent component of individual productivity z_t was set to match the Gini coefficient of earnings in the general population obtained from the same study mentioned above. The resulting value of ϱ is 0.95 which is very close to the numbers most commonly found in the literature.

Finally I need to describe the choice of the tax and social security parameters. Government consumption was chosen to match the ratio of federal, state and local consumption to GDP reported in the Economic Report of the President (2005, Table B–8 and B–20) and computed as an average from 1959 to 2005. The resulting number is 0.192. The income tax rate τ was then set so that once estate taxation is kept into account the government budget balances. Estate taxation is characterized by two parameters: the exemption level and the marginal tax rate. The exemption level in the US economy is 600000\$ and average earnings were approximately 42000\$ in 1998, so the exemption level in the model economy was set at 15 times average earnings. Given the exemption level the marginal tax rate on estates was set so as to match a ratio of estate tax revenues to GDP of 0.2%, the same reported in Castañeda et al. (2002). As far as the design of the social security system is concerned the following procedure was adopted: according to Huggett and Ventura (2000), over the period 1990–1994 the hospital and medical payment per retiree averaged 7.72 % and 4.70% of US GDP per person over age 20. Thus, I set the fixed component of retirement benefits \bar{b} equal to $0.1242Y$ where Y is GDP per capita in the model economy. The pension component in the US social security system is obtained in the fol-

lowing way: first earnings are indexed to make them comparable across years, then an average of the 35 best realizations is computed. Then a progressive formula is applied to this average or AIME (Average Indexed Monthly Earnings) to obtain the pension benefit to which the worker is entitled. The marginal benefit schedule equaled 90% of AIME up to 0.2 times average earnings, 32% of AIME between 0.2 and 1.24 times average earnings and 15% of AIME above that threshold; finally there is an upper bound of 2.47 of AIME above which no further benefits can be claimed.¹² In calibrating the variable part $b^v(\theta)$ of the model economy pension benefits we observe that the inherited component of earnings ability θ is an approximate measure of relative average earnings of different types of agents and therefore we apply to it the benefit formula of the US economy. This amounts to collapsing the distribution of lifetime earnings of a given type of household to its mean. While this is a simplification it still represents a substantial improvement compared to the standard practice of having a fixed benefit for the whole population. The reason is that a fixed benefit with the same aggregate expenditure underestimates the replacement ratio of high income people and overestimates that of low income people thus inflating the difference in wealth accumulation related to saving for retirement. The current formulation mitigates this problem because each earning type enjoys a different pension benefit; consequently the average deviation of replacement rates from the correct one is reduced. The resulting share of pension expenditures in GDP in the model is 6.71% and the payroll tax τ_s needed to finance the two components of social security benefits is 14.33%, only one percentage point below the figure reported by Conesa and Krueger (1999) for the US economy. This suggests that the formulation of the social security system proposed here, even though simplified, reproduces quite well the main features of the US system. Mortality rates refer to the US male population and are taken from the “Berkeley Mortality Database”; the implied dependency ratio is 0.27 a figure that is somewhat higher than the 0.22 of the US economy.¹³ Table ?? reports the values of the parameters that are fixed across all the experiments.

6 Results

In this section the results of the quantitative analysis of the model are discussed. In the first subsection I report the results obtained using two different versions of the increasing return function; in the second one I summarize those results and present a discussion of the relation between the current formulation of increasing

¹²See Huggett and Ventura (2000) or Social Security Online (2004).

¹³This discrepancy is not surprising since we study steady states while the US are clearly not on a stationary demographic path.

Figure 1: Return Schedules

returns and the one in Quadrini (2000). The first of the two return functions is obtained by approximating the empirical return schedule derived from the 1998 issue of the SCF with a logistic function, while the second is obtained from the 1995 issue of the SCF using a hyperbolic function for the interpolation. The reason for this choice can be understood by looking at figures 1 and 2. In Figure 1 the continuous line represents the return schedule from the 1998 issue of the SCF and the dashed line represents the one computed using the 1995 issue. A look at the figure reveals that the return schedule computed using 1995 data lies below the one computed on 1998 data and the amount of the difference is larger at higher levels of wealth. As a result the difference in the average portfolio return between the wealthiest and poorest households is slightly less than 4 percentage points in 1995 and is almost 5 percentage points in 1998. In Figure 2 the continuous lines depict the empirical schedules, while the dashed lines depict the interpolating functions. If we look at the bottom panel of figure 2 we see that the hyperbolic return function follows its empirical counterpart perfectly at low to medium levels of wealth but then underestimates it. A look at the top panel of figure 2 reveals that the logistic approximation, while still very good, somewhat overstates returns at medium-to-high wealth levels and then converges to the empirical values better as wealth further increases. For these reasons we may take results using the logistic approximation to 1998 data as an upper bound on the magnitude of the effects of increasing returns to savings

Figure 2: Empirical and interpolated return functions

and the hyperbolic approximation to 1995 data as a lower bound.

6.1 Steady State Distributions

The results presented here are the outcome of the following experiments: for each given return function the model is solved first with increasing returns, then the traditional case of constant interest rate is considered. In the first run of the model the discount factor on the descendants' utility γ is set so that the ratio of median bequests to average earnings is equal to its empirical counterpart. When subsequently solving the model with constant interest rate the subjective discount factor and the tax rates are adjusted so that the market clearing condition and the government budget constraint are still satisfied but the discount factor applied to the descendants' utility is kept constant. The reason is that in this way it is possible to assess the marginal effect of increasing returns to savings in the model without mixing it with changes in intergenerational altruism. The results for the logistic return function are reported first. The value of γ and β are 0.32 and 1.003. With these values the model generates a bequest-to-earnings ratio of 1.06 that matches very precisely the target and generates a capital-output ratio of 3.04 which equates the target of 3. When the model is run with constant returns agents need to be more impatient for

Table 5: The Distribution of Earnings and Wealth

Percentage share of Earnings by quantiles						
	Gini	Bottom 40	Top 20	Top 10	Top 5	Top 1
Data	0.63	2.8	61.4	43.5	31.1	14.8
Model	0.63	3.6	64.1	42.8	28.4	8.3
Percentage share of Wealth by quantiles						
	Gini	Bottom 40	Top 20	Top 10	Top 5	Top 1
Data	0.78	1.35	79.5	66.1	53.5	29.5
Model (L)	0.79	1.1	84.9	68.0	49.6	17.3
Model (C)	0.67	3.7	70.5	50.1	33.4	10.9

market clearing to hold at a capital-output ratio of 3 so that β is now 0.989; as γ is kept constant the bequest-to-earnings ratio changes to 2.15, double the one observed in the data. Recall that the ratio involves median bequests, so this increase translates into an equalizing effect on the wealth distribution. Summary statistics for the steady-state distributions are reported in Figure 3 and Table 5. In the top panel of Table 5 I report the distribution of earnings in the data and in the model. This distribution is exogenous to the model, however it is useful to assess its ability to match the empirical one. Recall from the calibration section that the parameters governing the earnings process were chosen so as to match the overall and first-year earners Gini index. The first of this calibration goal shows up in that the Gini index of the model and data reported in the table is the same. More importantly the calibrated earnings process can reproduce its empirical counterpart virtually perfectly up to the top 5 percent of the distribution and only attributes a slightly lower share of earnings to the top 1 percent. The latter is 14.8 percent in the data and 8.3 in the model. The ability of the earnings process used here to match earnings inequality in the data is then very satisfactory; even at the top of the distribution the difference is only a few percentage points, moreover this difference can be thought of as partially offsetting the omission of progressive taxation from the model. As far as the wealth distribution is concerned, Figure 3 gives a visual representation of the effects of increasing returns to savings. It is apparent that the small difference in returns faced by agents with different levels of assets is sufficient to generate a substantial amount of extra wealth inequality compared to the one implied by the assumption that all agents in the economy face the same returns on asset holdings. This can be seen by comparing the dotted line that represents the Lorenz curve in the model with constant return to savings and the dashed line representing the Lorenz curve in the model with increasing returns to savings: while the former always lies above the Lorenz curve drawn using the data —

Figure 3: Lorenz Curves: Logistic

depicted by the continuous line — the latter follows it quite closely. The bottom panel of Table 5 makes this point clearer. The Gini index of wealth in the model with constant interest rate is only 0.67 while the one of the model with increasing returns is 0.79. This value exactly matches the one computed for the US economy. A closer look at the wealth distribution reveals that the model with increasing returns fares almost perfectly up to the top 5 percent of the population and shows some weakness at the top 1 percent of the distribution attributing to it about 10 percentage points less of total net worth than the data. For comparison the model with constant interest rate underestimates the share of the top percentiles of the wealth distribution by a large amount quantifiable at 20 percentage points for the top 5 and top 1 wealthiest households; on the contrary it attributes to the bottom 40 percent of the population 3 times the share of wealth we observe in the data.

Next I report the results that are obtained when the chosen return schedule is based on the 1995 SCF and the hyperbolic interpolating function of equation 12 is used to model increasing returns to savings: again the parameters of the function are chosen so as to minimize the sum of squared deviations of the interpolating function from the empirical data and the discount on the descendants' utility is set so as to match the ratio of median bequests to average earnings. The value of γ that achieves that is 0.25. As a consequence of the heavier discounting of descendants' utility the value of β needed for market clearing climbs

Figure 4: Lorenz Curves: Hyperbolic

Table 6: Wealth Distribution: Hyperbolic Return Function

	Percentage share of Wealth by quantiles					
	Gini	Bottom 40	Top 20	Top 10	Top 5	Top 1
Data	0.78	1.35	79.5	66.1	53.5	29.5
Model (L)	0.76	1.3	81.0	60.6	41.3	13.6
Model (C)	0.68	3.2	71.7	51.5	34.4	11.1

up to 1.013. The median bequest-to earnings ratio is 1.017 and capital-output ratio from the supply side is 3.016 confirming the excellent quality of the solution to the market clearing and calibration system of equations. Results from this case are reported in Figure 4 and Table 6.¹⁴ Once again the lorenz curve of the constant interest rate case lies above the one of the increasing returns and then further from the data curve but differences are now smaller. This can be better understood by looking at the corresponding table. As it can be seen by comparing the second and third row of the table the Gini index goes up from 0.68 to 0.76 and the share of all the top percentiles of the wealth distribution increases substantially except for the top 1 percent of the population where extra wealth accumulation brought about by the higher return is negligible. Correctly the model with increasing return to savings cut by more than half the share of aggregate wealth held by the bottom 40 percent of the population. If we compare the model with hyperbolic returns to the data we see that the model can match perfectly the Gini index of wealth and the share of up to the top 10 percent of the population; it still does well at the top 5 percent of the distribution but the share of the top 1 percent is under-predicted by a large amount.

6.2 Summary and Discussion

In this section I summarize the results of the previous section first, then I discuss the relation between the current model with the closest approach in the literature, that is, the one based on entrepreneurship presented by Quadrini (2000). Table 7 reports the figures that are needed to develop the first point. In the first line of the table I report once more the wealth inequality statistics from the data. The second line performs a counterfactual experiment with those data, that is, it reports what the same distribution would be if the top 0.5 percent

¹⁴The nature of the experiment is the same as the one in the previous section: first the model with increasing returns to savings is solved, then this potential source of wealth inequality is shut off and the model is solved again. In particular as before, the intergenerational discount factor is kept constant across the two experiments; notice that since γ has changed across the two experiments the wealth distribution in the case of constant returns will show some minor changes here as well

held the same average net worth as the next half percentile, the one between the 99.0 and the 99.5 percent wealthiest households. The third line reports the average of the two models' steady state distribution obtained in the previous section. This distribution is not the result of any particular run of the model: as it was said before the two different increasing return functions deliver an upper and lower bound of the wealth inequality that results when we introduce in a model the fact that investment opportunities are related to wealth. The best assessment of this force will lie somewhere in between those bounds so reporting the average of the two distributions is a way to provide the reader a quick access to it. Finally the last line reports the steady state distribution of a model with constant interest rate when γ is calibrated so that the model generated bequest-to-earnings ratio is 1 as in the data, so it can be considered the "best" simulation for the constant return case. If we first compare the two model distributions to the true data distribution we see that increasing returns to savings do help getting a better match between the two: the Gini index in this case is exactly equal to its data counterpart while it is 9 points less in the constant interest case. Moreover the distribution in the increasing returns world adheres quite well to the one in the data and only departs from it substantially at the top 1 percent where it attributes only 50 percent of the true share. On the contrary the model with constant returns shows great difficulties at matching the data starting from the top 5 percent of the distribution: it gives a share that is 20 percentage points below the true one to the top 5 percent and about a third of the true value to the top 1 percent. When we compare the models to the data that do not include the extra net worth accumulated by the top 0.5 percent we see that the fit of both improves. However while the model with constant returns still attributes the top 5 and 1 percent of the distribution only about two thirds of the shares of wealth in the modified data, the model with increasing returns matches the distribution that excludes the top 0.5 percent extra wealth virtually perfectly. This leads us to the following two conclusions. First the fact that investment opportunities improve with wealth levels does add something to our understanding of the wealth distribution since it allows us to explain the accumulation behavior of households up to the top 0.5 percent wealthiest families, while this is not what happens in a similar model with constant returns. Second, the fact that the top 0.5 of the distribution is left unaccounted suggests that the simple formulation of the properties of returns used here does not change the view, well expressed for example in Carroll (2000) that there is something special about wealth accumulation by the very rich. Quadri's explanation based on entrepreneurship is a natural candidate in this respect since even within the top 1 percent of the population the share of entrepreneurs increases as we move further rightward in the tail of the distribu-

Table 7: Wealth Distribution: Summary

	Percentage share of Wealth by quantiles					
	Gini	Bottom 40	Top 20	Top 10	Top 5	Top 1
Data	0.78	1.35	79.5	66.1	53.5	29.5
Data (99.5%)	0.77	1.6	75.8	60.0	45.1	16.8
Model (L/H)	0.78	1.2	83.0	64.3	45.5	15.5
Model (C)	0.69	2.6	72.6	51.8	33.4	11.2

tion: in the top 1 percent of the wealth distribution 65 percent of households are entrepreneurs, this fraction climbs up to 75 percent in the top 0.5 percent and to 82 percent in the top 0.25 percent.¹⁵ For this reason it is worth comparing the present formulation of increasing returns to savings to the approach based on entrepreneurship. In practice Quadrini models entrepreneurship as a set of projects of increasing size that an agent may have the chance to undertake. In so doing the agent would add business profits to regular wage earnings. For each project size there is a range of wealth where the agent may undertake the project but has to finance it partially with loans at a rate that is above market rates. This specification is equivalent to a return function that in some ranges of wealth exhibits returns that are above the market interest. This is in part because any dollar invested raises the probability that the agent will be able to take advantage of the extra profits from running a business —or one that is larger than the one he is already running—, in part because as long as he needs to finance the business through loaned funds, the marginal return to savings will sum the intermediation cost to the market return. Otherwise the model is a standard infinite-horizon buffer stock model. This latter fact implies that particular combinations of patience and returns to investment in the firm would make the agent invest any amount of wealth. Indeed in his calibration the condition for boundedness of asset holdings is not met for entrepreneurs that are borrowing. However as the maximum firm size — the largest of the three possible projects — is reached returns fall again to market levels and accumulation rapidly stops. By carefully choosing the size of entrepreneurial projects and therefore the implicit return function Quadrini can reproduce the wealth distribution. My approach is simpler: I just estimate a return function from data and assess its quantitative impact on the wealth distribution. On the one hand this approach is more general since the fact that wealthier agents face higher returns through different investment choices is not confined to entrepreneurs. On the other hand it misses some peculiar features of entrepreneurial activity like the

¹⁵These figures are based on the 1998 issue of the SCF and define an entrepreneurial household one that has non-zero investment in closely held businesses.

role played by financial imperfections to increase the cost of external financing or the existence of idiosyncratic investment risk.¹⁶ The other main difference between the current framework and the one in Quadrini is the assumption made here that households have finite lives and form dynasties through ability and financial links. In a finite horizon model the mechanism that lies at the hearth of Quadrini’s paper, that is, the combination of patience and entrepreneurial returns would not hold any more. It would then be interesting to see if and under what conditions bequests, life-cycle savings and other issues that arise in this context would allow the results of the entrepreneurial model to survive. This is beyond the scope of the present paper and is left for future research.

7 Conclusions

The economic fortunes of the households in real economies are very unequal with wealth being substantially more concentrated than other measures like income and earnings. This fact has attracted a lot of attention among macroeconomists. The basic framework of precautionary saving outlined in Aiyagari (1994) has proven capable of reproducing the fact reported above qualitatively, but not quantitatively leading to successive extensions that include different features, like social security, intergenerational links, entrepreneurship or heterogeneous preferences but retain the basic assumption that a single asset is available in the economy. This paper explicitly acknowledges the fact that in reality there is a menu of assets with different returns that households may use to carry out their saving plans and that there is a systematic positive relationship between asset holdings and the return to these holdings.

To accomplish this task I have considered a model that blends the life-cycle and dynastic framework where I assume that agents face a return to their savings that is increasing in the level of assets they hold. This feature is able to increase substantially the level of wealth inequality compared to the standard case of constant returns. As a matter of fact the model is able to account for the U.S. distribution of wealth up to the 99.5 percentile, whereas a similarly calibrated model that omits the relationship between asset holdings and returns cannot. The model still fails to match the huge fortunes that are accumulated by a few wealthy households at the very top — 0.5 percent — of the distribution. This supports the view that some other mechanism must be at work for this group. Entrepreneurship as proposed by Quadrini or the the capitalist spirit as proposed by Carroll are two possibilities. Integrating those two theories in a

¹⁶Quadrini also assumes that entrepreneurial activity carries extra income risk magnifying precautionary savings of business households compared to workers. This aspect is not central to the comparison with the model presented in this paper so I won’t elaborate it further

full life-cycle models with bequest like the one proposed here is left for future research.

A Data Construction

In this appendix I describe briefly the construction of the household portfolios from SCF data. Family asset holdings were classified into eight categories: liquid accounts, bonds, stocks, other financial assets, primary residence, investment real estate, business equity and other non financial assets. Liabilities were classified into four broad groups: mortgage and other loans on primary residence, other property loans, credit card balances and a residual category that includes other type of debt. The purpose of this classification is to reduce the complexity of household portfolios to a small number of asset types to which we can assign return data. The classification proposed sometimes does not overlap perfectly with the one in the SCF thus requiring some imputations. This is the case of defined contribution pension plans, IRA and Keogh accounts and trusts.¹⁷ In these cases the SCF reports qualitative information about how these accounts were invested. An example of the list of answers proposed to the respondent to the SCF questionnaire is: mostly stock, mostly bonds, a combination of the two or a menu of other types of assets. In this case if the answer was mostly stocks or mostly bonds I considered the whole account invested in stock or bond, if the answer was a combination, the imputation was half to stocks and half to bonds, in the other cases the account was included in the residual category of the classification. Summarizing the exact description of the single items in the classification used here is:

- Liquid accounts: checking and savings accounts, money market mutual funds, cash call accounts at brokers, certificates of deposits plus all of the previous items held through retirement accounts, trusts and other managed accounts.
- Bonds: all local, state, federal, corporate and foreign bonds held directly, through mutual funds or through retirement accounts, trusts and other managed accounts.
- Stock: all stock owned directly or through mutual funds, retirement accounts, trusts and other managed accounts.
- Other financial assets: this category includes a broad set of assets ranging from cash value of life insurance policies, to loans to friends. It also

¹⁷Except for the imputations, the classification follows the one in Bertaut and Starr-McCluer (2000) to which the reader is referred for more details than those reported in this appendix

includes the share of retirement accounts, trusts and other managed accounts invested in assets other than liquid, bonds and stocks of for which no investment answer was reported. Finally it includes mutual funds not classified.

- Primary residence: value of home, mobile home plus land and site where it stands.
- Investment real estate: value of all other property.
- Business: value of equity in closely held firms of any legal form
- Other nonfinancial assets: vehicles, jewelry, artwork, airplanes and a large number of other possible real assets.

Once the information on asset and liabilities holdings was classified, total net worth was computed and the households were ordered based on it. For each of the net worth percentile group defined in the main text the share of each category of asset in total net worth and total wealth was computed. These data formed the basis on which the empirical return function was constructed.

B Numerical Solution Method

The method used to solve the model are quite standard and the interested reader is referred to Ríos-Rull (1998) for a detailed account. What I will do in this section is first to give a short outline of the procedure used and then to describe in more detail the method used to solve the dynasty dynamic programming problem. As far as the first part is concerned the steps are the following:

1. Guess initial values of the intergenerational discount factor γ and the tax rate on bequests τ_e .
2. Guess values for the subjective discount factor β and the income tax rate τ .
3. Solve the dynasty optimization problem and obtain the decision rules for assets and consumption. The procedure followed implied computing the value function at each age and finding the decision rules by maximizing the right-hand side of the Bellman equation rather than finding the zero of the associate euler equation. More details are given below.
4. Use the decision rules from step 3 and random sequences from the survival and labor endowment processes to simulate the economy, obtain aggregate asset holdings and tax revenues; if the market clearing and government budget balance conditions are met move on, else go back to 2.

5. If the bequest-to-earnings and the share of estate tax revenue in GDP targets are met then move on, else go back to 1.
6. Compute the statistics of interest on the steady state distribution.

For the simulation of the economy the history of 40000 dynasties for 3000 periods were used; further increasing those numbers did not affect the results. Given that a simple condition to insure that asset holdings are bounded above could not be found particular care was taken to avoid that the results could be an artefact of the finite simulation horizon. In practice I kept track of the maximum realized value of asset holdings in each period and checked that it did not trend upward in the simulation.

Next I describe the two issues related to the solution of the dynastic optimization problem. First notice that under the assumption of pure altruism an agent cares about its descendants's utility, however the value function of the latter is not known at the start of the algorithm. The general structure is then that of a fixed point problem: starting from an initial guess for the age 1 value function V_1^0 , the household decision problem is solved. This delivers a set of age indexed value functions including a new guess for the age 1 value function, say V_1^1 . Again the household decision problem is solved and then the procedure is repeated until convergence in the sequence V_1^n is obtained up to a precision of 10^{-6} .

The second issue here is that the increasing return to asset assumption introduces non-concavities in the value function giving rise to the possibility of multiple local solutions. This problem is compounded by the fact that in a large range of wealth the difference in returns between the two cases of constant and increasing returns may be minor translating into minor differences in the decision rules: the algorithm used should then be sufficiently accurate to distinguish between the two. To address the first concern a direct search algorithm was used. In practice let $S = \{s_1, \dots, s_i, \dots, s_n\}$ be the set of points used to discretize asset holdings. For each interval $[s_i, s_{i+1}]$ ten equally spaced points were laid; the resulting set of points defines the action space. The value function at points of the action space that do not coincide with state space points was evaluated by using cubic spline approximation. The optimal solution was found by searching the whole set of possible actions defined by the intersection of the finer grid and the individual budget constraint. This search gives a first approximation to the solution, say a^* . The next step is meant to address the second concern. It consists of defining a symmetric interval around the first approximate solution. This interval has radius d defined as:

$$d = \begin{cases} a^*/10 & \text{if } a^*/10 < 1 \\ \sqrt{a^*/10} & \text{if } a^*/10 > 1 \end{cases}$$

In this interval 200 equally spaced points were laid and a second direct search over all these points was performed to find the solution a^{opt} . This double-search procedure guarantees a high degree of accuracy in the solution with only a few hundreds function evaluations rather than the many thousands that would be needed if a single step search was performed. The method was tested on periodic functions like $f(x) = x \sin x$ and invariably found the correct absolute maximum — with a high degree of precision — among the many local solutions that such functions have.

References

- [1] Aiyagari, S.R. (1994). “Uninsured Idiosyncratic Risk and Aggregate Saving”. *Quarterly Journal of Economics*, 109(3): 659-684.
- [2] Barsky R. B., Juster F. T., Kimball M. S. and Shapiro M.D. (1997). “Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study”. *Quarterly Journal of Economics*, 112 (2): 537-579.
- [3] Bertaut C. and Starr-McCluer M. (2000). “Household Portfolios in the United States”. *Federal Reserve Board of Governors, Finance and Economics Discussion Paper*: 26 (April).
- [4] Brent R. (1973). “Algorithms for Minimization without Derivatives”. Englewood Cliffs, NJ: Prentice-Hall.
- [5] Budría-Rodríguez S., Díaz-Giménez J., Quadrini V. and Ríos-Rull J.-V. (2002). “Updated Facts on the U.S. Distributions of Earnings, Income and Wealth”. *Federal Reserve Bank of Minneapolis Quarterly Review*, 26, (Summer): 2-35.
- [6] Cagetti M. and De Nardi M. (2002). “Entrepreneurship, Frictions and Wealth”. *Federal Reserve Bank of Minneapolis Working Paper* 620.
- [7] Carroll C.D. (2000). “Why Do the Rich Save So Much ?”. In: J. B. Slemrod, (Ed). “Does Atlas Shrug? The Economic Consequences of Taxing the Rich”. Cambridge, MA: Harvard University Press,
- [8] Castañeda A., Díaz-Giménez J. and Ríos-Rull J.-V. (2002). “Accounting for the U.S. Earnings and Wealth Inequality”. Forthcoming *Journal of Political Economy*.
- [9] Cocco J., Gomes F. and Maenhout P. (2005). “Consumption and Portfolio Choice over the Life Cycle”. *Review of Financial Studies*. 18(2): 491-533.

- [10] Conesa M. and Krueger D. (1999). "Social Security Reform with Heterogeneous Agents". *Review of Economic Dynamics*, 2: 757-795.
- [11] De Nardi M. (2001). "Wealth Inequality and Intergenerational Links". Forthcoming *Review of Economic Studies*.
- [12] Díaz-Giménez J., Quadrini V. and Ríos-Rull J.-V. (1997), "Dimensions of Inequality: Facts on the U.S. Distribution of Earnings, Income and Wealth". *Federal Reserve Bank of Minneapolis Quarterly Review*, 21, (Spring): 3-21.
- [13] Domeij D. and Klein P. (2002). "Public Pensions: To What Extent Do They Account for Swedish Wealth Inequality ?". *Review of Economic Dynamics*, 5: 503-534.
- [14] Gentry W. M. and Hubbard G.R. (2001). "Entrepreneurship and Household Saving". *National Bureau of Economic Research, Working Paper* No. 7894.
- [15] Goetzmann W. N. and Spiegel M. (2000). "The Policy Implications of Portfolio Choice in Underserved Mortgage Markets". *Yale International Center for Finance. Working Paper* 00-18.
- [16] Guiso L., Haliassos M. and Jappelli T. eds. (2002). "Household Portfolios". Cambridge and London: MIT press.
- [17] Guiso L., Sapienza P. and Zingales L. (2004). "The Role of Social Capital in Financial Development". *American Economic Review*. 94: 526-556.
- [18] Hansen G. (1993). "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked". *Journal of Applied Econometrics*. 8(1): 71-80.
- [19] Hong H., Kubik J. and Stein J. (2001). "Social Interaction and Stock Market Participation". *National Bureau of Economic Research, Working Paper* No. 8358.
- [20] Hubbard G. R., Skinner J. and Zeldes S. (1995). "Precautionary Saving and Social Insurance". *Journal of Political Economy*. 103(2): 360-399.
- [21] Huggett M. (1996). "Wealth Distribution in Life-Cycle Economies". *Journal of Monetary Economics*, 38(3):469-494.
- [22] Huggett M. and Ventura G. (2000). "Understanding Why High Income Households Save More than Low Income Households". *Journal of Monetary Economics*, 45: 361-397.

- [23] Hurd M. and Smith J. (2002). "Expected Bequests and Their Distribution". *National Bureau of Economic Research, Working Paper* No. 9142.
- [24] Hurst E., Luoh M.C. and Stafford F.P. (1998), "The Wealth Dynamics of American Families, 1984-94". *Brookings Papers on Economic Activity*, 1: 267-329.
- [25] Jagannathan R., McGrattan E. and Scherbina A. (2000). "The Declining U.S. Equity Premium". *Federal Reserve Bank of Minneapolis Quarterly Review*, 24 (Fall): 3-19.
- [26] Kennickell A. B., Starr-McCluer M. and Surette B.J. (2000). "Recent Changes in U.S. Family Finances: Results from the 1998 Survey of Consumer Finances". *Federal Reserve Bulletin*, (January): 1-29.
- [27] Krusell P. and Smith A.A. (1998). "Income and Wealth Heterogeneity in the Macroeconomy". *Journal of Political Economy*, 106(5): 867-896.
- [28] Moskowitz T. and Vissing-Jørgensen A. (2002). "The Returns to Entrepreneurial Investment: A Private Equity Premium Puzzle ?". *American Economic Review*, 92(4): 745-778.
- [29] Paiella P. (2001). "Limited Financial Market Participation: A Transaction-Cost Based Explanation". The Institute for Fiscal Studies. WP 01/06.
- [30] Quadrini V. (2000). "Entrepreneurship, Saving and Social Mobility". *Review of Economic Dynamics*, 3,(1): 1-40.
- [31] Ríos Rull J.V. (1999). "Computation of Equilibria in Heterogeneous-Agent Economies ". In: Marimon R. and Scott A. Eds. "Computational Methods for the Study of Dynamic Economies". Oxford and New York: Oxford University Press.
- [32] Samwick A.A. (2000). "Portfolio Responses to Taxation: Evidence from the End of the Rainbow". In: J. B. Slemrod, (Ed). "Does Atlas Shrug? The Economic Consequences of Taxing the Rich". Cambridge, MA: Harvard University Press.
- [33] Vissing-Jørgensen A. (2001). "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures". Working Paper. University of Chicago.
- [34] Edward N. Wolff. (2000). "Who Are the Rich ? A Demographic Profile of High-Income and High-Wealth Americans". In: J. B. Slemrod, (Ed). "Does Atlas Shrug? The Economic Consequences of Taxing the Rich". Cambridge, MA: Harvard University Press.